

The Study of Semantic and Ontological Features of Thesaurus and Ontology-based Information Retrieval Systems

Elmira Karimi*

PhD Candidate in Library and Information Science; Iranian Research Institute for Information Science and Technology (IranDoc) Email: karimi@students.irandoc.ac.ir

Mahmoud Babaei

Assistant Professor in Iranian Research Institute for Information Science and Technology (IranDoc) Email: babaie@irandoc.ac.ir

MolukoSadat Hosseini Beheshti

Assistant Professor in Iranian Research Institute for Information Science and Technology (IranDoc) Email: beheshti@irandoc.ac.ir

Received: 22, Apr. 2016

Accepted: 17, Apr. 2018

Abstract: In order to increase relevant retrieved results and users satisfaction, information retrieval systems are seeking a way to change keyword-based search to concept-based search. These systems use the tools that can convert terms to concepts. Over the years, this provided the basis for production of tools such as classification schema, vocabularies, subject headings and thesauri. Recently, ontologies are used in complementing this process.

The main problem of this research is how the great information retrieval systems of the world claiming to use ontologies and thesauri use the semantic features of these tools in their information retrieval process? In this paper 7 information retrieval systems which use these tools, were selected. Using literature review, the semantic features were presented in the form of research checklist. These features include 5 general categories of concept search, synonyms search, suggestion for search, use of ontologies and display. The researcher examined these features with structured observation in the form of checklist. Though these systems use these tools for searching and retrieving, they don't show the expected features. In studied factors, ERIC Database and the National Library of Australia achieved maximum scores of 32 and 24 respectively, while the National Archive of Australia and the British Museum had minimum scores of 12 and 13 respectively.

Keywords: Semantic Search, Semantic Features of Information Retrieval, Information Retrieval Systems, Concept Search, Synonyms Search, Ontology

* Corresponding Author

**Iranian Journal of
Information
Processing and
Management**

**Iranian Research Institute
for Information Science and Technology
(IranDoc)**

ISSN 2251-8223

eISSN 2251-8231

Indexed by SCOPUS, ISC, & LISTA

Vol. 34 | No. 4 | pp. ?00-?00

Summer 2019



بررسی ویژگی‌های معنایی و هستی‌شناسانه نظام‌های بازیابی اطلاعات مبتنی بر اصطلاح‌نامه و هستی‌شناسی

المیرا کریمی

دانشجوی دکتری علم اطلاعات و دانش‌شناسی؛
پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)؛
پدیده‌آور رابط karimi@students.irandoc.ac.ir

محمود بابایی

دکتری علوم ارتباطات؛ استادیار؛
پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)؛
babaie@irandoc.ac.ir

ملوک‌السادات حسینی بهشتی

دکتری زبان‌شناسی؛ استادیار؛ پژوهشگاه علوم و فناوری
اطلاعات ایران (ایرانداک) beheshti@irandoc.ac.ir



دریافت: ۱۳۹۵/۱۰/۰۷ | پذیرش: ۱۳۹۶/۰۳/۲۴ | مقاله برای اصلاح به مدت ۱۶ روز نزد پدیدآوران بوده است.

فصلنامه | علمی پژوهشی
پژوهشگاه علوم و فناوری اطلاعات ایران
(ایرانداک)

شاپا (چاپی) ۸۲۳۳-۲۲۵۱

شاپا (الکترونیکی) ۸۲۳۱-۲۲۵۱

نمایه در SCOPUS، ISI، LISTA و

jipm.irandoc.ac.ir

دوره ۳۴ | شماره ۴ | صص ۱۵۷۹-۱۶۰۶

تابستان ۱۳۹۸

چکیده: نظام‌های بازیابی اطلاعات برای افزایش نتایج مرتبط بازیابی شده و کسب رضایت کاربران به دنبال تغییر از جست‌وجوی مبتنی بر کلیدواژه به جست‌وجوی مبتنی بر مفاهیم هستند. این نظام‌ها برای انجام این کار از ابزارهایی استفاده می‌کنند که بتوانند واژگان را به مفاهیم تبدیل کنند. این امر طی سالیان متمادی زمینه تولید ابزارهایی همچون طرح‌های رده‌بندی، سرعنوان‌های موضوعی، واژگان‌ها و اصطلاح‌نامه‌ها را فراهم آورده است. در تکامل این روند، هستی‌شناسی‌ها نیز به تازگی به کار گرفته شده‌اند. چگونگی بهره‌گیری از ویژگی‌های معنایی ابزارهایی مانند هستی‌شناسی و اصطلاح‌نامه در روند بازیابی اطلاعات برای نظام‌های مطرح دنیا که مدعی بازیابی اطلاعات با استفاده از این ابزارها هستند، مشکل اصلی مورد بررسی در این پژوهش است.

در این مقاله هفت نظام بازیابی اطلاعات با توجه به استفاده از این ابزارها انتخاب شدند و با بررسی متون، ویژگی‌های معنایی نظام‌ها در قالب سیاهه و ارسلی محقق‌ساخته‌ای ارائه شد. این ویژگی‌ها پنج مقوله کلی جست‌وجوی مفهومی، جست‌وجوی مترادف‌ها، پیشنهاد برای جست‌وجو، استفاده از هستی‌شناسی‌ها و نمایش هستند. پژوهشگر با مشاهده ساختار یافته در قالب سیاهه، این ویژگی‌ها را مورد بررسی قرار داد. با وجود این که این



نظام‌ها از این ابزارها برای جست‌وجو و بازیابی اطلاعات بهره گرفته‌اند، اما ویژگی‌های معنایی و هستی‌شناسانه مورد انتظار را نشان ندادند. «پایگاه اطلاعاتی اریک» و «کتابخانه ملی استرالیا» بیشترین امتیاز (۳۲ و ۲۴) را در معیارهای بررسی‌شده به‌دست آوردند و «آرشیو ملی استرالیا» و «موزه بریتانیا» کمترین امتیاز (۱۲ و ۱۳) را داشتند.

کلیدواژه‌ها: جست‌وجوی معنایی، ویژگی‌های معنایی بازیابی اطلاعات، نظام‌های بازیابی اطلاعات، جست‌وجوی مفهومی، جست‌وجوی مترادف‌ها، هستی‌شناسی

۱. مقدمه و مسئله

جست‌وجوی اطلاعات، فعالیت انسانی پیچیده‌ای است که با یک نیاز اطلاعاتی آغاز می‌شود و در بعضی موارد، چندین شکل استراتژی جست‌وجو و مرور منابع اطلاعاتی گوناگون را شامل می‌شود. هر مدلی از جست‌وجوی اطلاعات شرحی از مفهوم نیاز اطلاعاتی را که فرایند جست‌وجو با آن شروع می‌شود، فراهم می‌کند. از دیدگاه شناختی، این نیازها که در بسیاری از موارد مبهم هستند، ممکن است از طریق مفاهیم سطح بالا بیان شوند (Sutcliffe and Ennis 1998) که به‌نوبه خود به مفاهیمی دیگر مرتبط بوده و بخشی از دامنه خاص یا درون تعداد کمی از آن‌ها هستند (García and Sicilia 2003). یافتن پاسخی برای این نیاز اطلاعاتی از سوی نظام‌ها، نیازمند دریافت معنای دقیق پرس‌وجوی کاربر است.

نظام‌های بازیابی اطلاعات تاکنون اغلب به کلمات و کلیدواژه‌ها بیش از معنای آن‌ها می‌پرداختند؛ به‌نوعی که در اکثر موارد معنای کلیدواژه مورد نظر کاربر نادیده گرفته می‌شد. این نظام‌ها از فهرست‌های کلیدواژه‌ای برای توصیف محتوای اطلاعاتی استفاده می‌کنند. اما مسئله این است که چنین فهرستی چیزی در مورد ارتباطات معنایی بین کلیدواژه‌ها نمی‌گوید و معنای کلمات و عبارات را مورد توجه قرار نمی‌دهد (Styltvisg 2006). همان‌طور که «رید» اشاره می‌کند:

«خطر بزرگ در پرداختن به کلمات این است که کلمه را به‌جای معنا در نظر می‌گیرند. این غلط است. یک کلمه، در چارچوب اجتماعی ایجاد می‌شود که می‌تواند با چند جنبه یا بعد که همگی در یک زمان ارائه می‌شوند، در نظر گرفته شود. توصیف معنایی کامل یک کلمه شامل حداقل چهار جنبه یا بعد است: «واجگان»، «نحو»، «ساخت‌واژه» و «کاربرد» (Read 1942).

اغلب، برای کاربران عادی استفاده از سیستم‌های اطلاعاتی بر پایه فنون کلیدواژه‌محور مشکل است. کاربران معمولاً در بیان نیازهای اطلاعاتی خود و ترجمه نیازها به سؤال با مشکل مواجه‌اند.

نوع رفتار جست‌وجوی کاربران مستلزم داشتن دانش جدید و فنون بازیابی و راهبردهای بازیابی جدید است. بازیابی اسناد از طریق شیوه‌های معنایی^۱، مبتنی بر تحلیل ساختارهای معنایی محتوای اسناد صورت می‌گیرد (Hynek 2002). فنون مختلفی بر اساس هوش مصنوعی و پردازش زبان طبیعی برای پردازش معنایی به کار می‌روند و بیشتر آن‌ها به استفاده از ساختارهای کمکی همچون واژگان کنترل‌شده و هستی‌شناسی‌ها متکی هستند. با استفاده از واژگان کنترل‌شده (واژه‌نامه‌ها و اصطلاح‌نامه‌ها) اصطلاحات اخص، اعم و مرتبط در پرس‌وجوها گنجانده می‌شوند. واژگان کنترل‌شده یکی از روش‌های غلبه بر بعضی از شدیدترین موانع پرس‌وجوهای کلیدواژه‌ای هستند. طی سال‌ها، ساختارهای کمکی دیگری نیز مثل مجموعه‌های بزرگ مترادف «وردنت»^۲ برای این هدف ایجاد شد. مشخص شده است که جست‌وجوی مفهومی بر اساس ساختارهای کمکی مثل «وردنت» می‌تواند به طوری کارآمدتر با استفاده مجدد از مدل‌های بازیابی اطلاعات و ساختارهای داده‌ای از بازیابی اطلاعات کلاسیک به کار رود. مدل‌های داده‌ای که مجموعه‌هایی از مفاهیم در دامنه خاص (هستی‌شناسی‌های دامنه‌ای) را ارائه می‌دهند، و آن‌هایی که می‌توانند ارتباطات میان اصطلاحات را دربرگیرند، در سال‌های اخیر به کار گرفته شده‌اند.

بنابراین، طبق آنچه گفته شد، مشخص است که مشکلات جست‌وجوی کلیدواژه‌ای و معنایی می‌تواند به بازیابی منابع غیرمرتبط با نیاز اطلاعاتی کاربر منجر شود. به این دلیل، لزوم توجه به معنا کاملاً احساس می‌شود. نظام‌های بازیابی بزرگ دنیا به دنبال یافتن راهی برای حل این مشکل هستند و هر یک از ابزارهای رفع این مشکل استفاده کرده‌اند. نظام‌های بازیابی اطلاعات موجود در ایران نیز به دنبال رسیدن به جست‌وجوی معنایی و حل مشکل جست‌وجوی کلیدواژه‌ای هستند. در ایران نیز تقریباً مانند بقیه دنیا، نظام جامعی که بتواند به ویژگی‌های معنایی بپردازد، وجود ندارد. بنابراین، این نظام‌ها در مرحله اول می‌توانند از رویکردهای مورد استفاده در نظام‌های بازیابی اطلاعات بزرگ

1. semantic methods

2. WordNet

دنیا ایده گرفته و سپس، با توجه به رویکردها و ویژگی‌های مورد استفاده آن‌ها، نظام بازتابی معنایی خود را پایه‌گذاری کنند.

این مقاله به دنبال پاسخ به این پرسش است که کدام ویژگی‌های معنایی در نظام‌های بازتابی اطلاعات که مدعی استفاده از ابزارهای معنایی هستند، مورد استفاده قرار گرفته‌اند و آیا با استفاده از این ابزارها به اهداف معنایی اشاره‌شده در متون رسیده‌اند یا خیر؟ برای بررسی این مسئله، محقق با بررسی متون به سیاهه‌ای دست یافت که بر اساس آن به مشاهده این نظام‌ها پرداخت.

۲. هدف‌های پژوهش

هدف اصلی این پژوهش بررسی عملکرد نظام‌های بازتابی اطلاعات مورد نظر پژوهش از نظر استفاده از ویژگی‌های معنایی است. برای رسیدن به این هدف، اهداف فرعی زیر دنبال خواهد شد:

۱. ارزیابی ویژگی‌های جست‌وجوی مفهومی در نظام‌های بازتابی اطلاعات مورد نظر پژوهش؛
۲. ارزیابی ویژگی‌های جست‌وجوی مترادف‌ها در نظام‌های بازتابی اطلاعات مورد نظر پژوهش؛
۳. ارزیابی ویژگی‌های پیشنهاد برای جست‌وجو در نظام‌های بازتابی اطلاعات مورد نظر پژوهش؛
۴. ارزیابی ویژگی‌های مربوط به استفاده از هستی‌شناسی‌ها در نظام‌های بازتابی اطلاعات مورد نظر پژوهش؛
۵. ارزیابی ویژگی‌های نمایش در نظام‌های بازتابی اطلاعات مورد نظر پژوهش.

۳. سؤال‌های پژوهش

سؤال اصلی پژوهش این است که عملکرد نظام‌های بازتابی اطلاعات مورد نظر پژوهش از نظر استفاده از ویژگی‌های معنایی چگونه است؟ برای پاسخ به این سؤال اصلی، سؤالات فرعی زیر دنبال خواهد شد:

۱. جست‌وجوی مفهومی در نظام‌های بازتابی اطلاعات مورد نظر پژوهش چگونه است؟

۲. ویژگی جست‌وجوی مترادف‌ها در نظام‌های بازیابی اطلاعات مورد نظر پژوهش چگونه است؟
۳. ویژگی پیشنهاد برای جست‌وجو در نظام‌های بازیابی اطلاعات مورد نظر پژوهش چگونه است؟
۴. ویژگی مربوط به استفاده از هستی‌شناسی‌ها در نظام‌های بازیابی اطلاعات مورد نظر پژوهش چگونه است؟
۵. ویژگی نمایش در نظام‌های بازیابی اطلاعات مورد نظر پژوهش چگونه است؟

۴. پیشینه پژوهش

مروری بر کارهایی که تاکنون در حوزه ارزیابی نظام‌های بازیابی اطلاعات انجام شده نشان داد که به‌طور کلی، تقریباً به هیچ‌یک از این نظام‌ها از دیدگاه معنایی نگاه نشده است و محققان اغلب برای ارزیابی این نظام‌های بازیابی اطلاعات رویکردهای کلی و تخصصی ارزیابی عملکرد، به‌جز ویژگی‌های معنایی، را در نظر گرفته‌اند. علاوه بر این، مرور متون نشان داد که بیشتر بررسی‌های انجام‌شده بر موتورهای جست‌وجوی معنایی متمرکز هستند. از میان بررسی‌های انجام‌شده بر روی موتورهای جست‌وجوی معنایی نیز تعداد کمی از آن‌ها به بررسی ساختاریافته پرداخته‌اند و اغلب مطالعات به مطالعه اجمالی و بررسی کلی این موتورهای جست‌وجو اکتفا نموده‌اند. در واقع، اغلب مطالعات با انجام جست‌وجوهای مختلف در این موتورهای جست‌وجو ویژگی‌های معنایی آن‌ها را به تصویر کشیده‌اند. (Turner, Shah and Bitirim (2009); Andago, Phoebe and Thanoun (2010); O' leary (2010; Sudeepthi, Anuradha and Babu (2012); Azizan et al. (2013); Negi and Kumar (2014); Kumar (2014) تماماً با انجام جست‌وجوهای به بررسی ویژگی‌های معنایی در موتورهای جست‌وجو پرداخته و اغلب آن‌ها فهرستی از امکانات و ویژگی‌های مورد استفاده در موتورهای جست‌وجو را ارائه کرده‌اند. در بعضی از آن‌ها از جمله در پژوهش «ترنر، شاه و بیتیریم» مقایسه‌ای بین موتورهای جست‌وجوی معنایی و موتورهای جست‌وجوی عادی انجام شده و در نهایت، به این نتیجه رسیده‌اند که عملکرد جست‌وجوی معنایی در هر دو گروه موتور جست‌وجوی کلیدواژه‌ای و موتور جست‌وجوی معنایی در سطح پایینی قرار دارد (Turner, Shah and Bitirim 2009). «آنداگو، فوبی و تانون» نیز این مقایسه را بین

«گوگل» و «هاکیا»^۱ انجام دادند. نتایج این پژوهش نشان داد که «گوگل» عملکرد بهتری نسبت به «هاکیا» دارد (Andago, Phoebe and Thanoun 2010). «اولری» نیز مانند «آنداگو»، فوبی و تانون» به مقایسه «گوگل» و «هاکیا» پرداخت و نتایج پژوهش او نیز عملکرد بهتر «گوگل» نسبت به «هاکیا» را نشان داد (O' leary (2010). «نگی و کومار» این مقایسه را بین دو موتور جست‌وجوی معمولی «گوگل» و «یاهو» و سه موتور جست‌وجوی معنایی «هاکیا»، «داک‌داک‌گو»^۲ و «بینگ»^۳ انجام دادند و در نهایت، به این نتیجه رسیدند که عملکرد جست‌وجوی معنایی موتورهای جست‌وجوی معنایی بالاتر از موتورهای جست‌وجوی عادی است (Negi and Kumar 2014).

پژوهش‌های «سودپتی، انورادها و بابو» و «عزیزان» و همکاران (۲۰۱۳) بر روی موتورهای جست‌وجوی معنایی انجام شد. هدف آن‌ها ارزیابی این موتورهای جست‌وجو نبود، بلکه آن‌ها به دنبال یافتن ویژگی‌ها و امکانات معنایی در این موتورهای جست‌وجو بودند و در نهایت، فهرستی از این ویژگی‌ها را ارائه کردند (Sudeepthi, Anuradha and Babu 2012; Azizan et al. 2013).

در ایران نیز «دری» در مقاله‌ای عملکرد پنج موتور جست‌وجوی معنایی را که به صورت رایگان در وب قابل دسترس هستند، با استفاده از ۴۵ معیار مشخص در قالب یک سیاهه و آرسی محقق ساخته مورد ارزیابی و مقایسه قرار داد. این سیاهه شامل دو بخش ویژگی‌های معنایی و عادی موتورهای جست‌وجو بود. نتایج پژوهش نشان داد که موتورهای جست‌وجوی مورد بررسی در شاخص ویژگی‌های عادی و معنایی، عملکرد مطلوب و کارایی مورد انتظار را نداشتند. نتایج مقایسه ویژگی‌های معنایی در این موتورهای جست‌وجو نشان داد که «داک‌داک‌گو» و «هاکیا» بیشترین امتیاز را در این مسئله دارا هستند (۱۳۹۳).

همان‌طور که در بالا اشاره شد، پژوهش‌های بسیار کمی به بررسی ویژگی‌های معنایی در نظام‌های بازیابی شناخته‌شده جهان پرداخته‌اند و از طرفی، از دیدگاه هستی‌شناسانه این ویژگی‌ها مورد توجه قرار نگرفته‌اند. این پژوهش با بررسی ویژگی‌های معنایی در نظام‌های بزرگ که از ابزارهای معنایی استفاده می‌کنند، به استفاده از هستی‌شناسی‌ها و

1. Hakya: <http://www.hakia.com/>

2. DuckDuckGo

3. Bing

ویژگی‌های مربوط به آن‌ها نگاه ویژه خواهد داشت.

۵. روش پژوهش

این پژوهش از نوع پژوهش‌های توصیفی-کاربردی است که به پیمایش نظام‌های بازیابی اطلاعات بزرگ دنیا می‌پردازد. در این پژوهش از مشاهده ساختاریافته با استفاده از سیاهه و آرسی محقق‌ساخته‌ای برای بررسی معیارها و دستیابی به داده‌های پژوهش استفاده شده است. برای تدوین این سیاهه، ویژگی‌های معنایی مورد نظر در پژوهش «دری» (۱۳۹۳) به‌عنوان اساس اولیه مورد توجه قرار گرفت که شامل ۴۵ معیار است. این سیاهه با مطالعه سایر منابع تکمیل گردید. روایی این سیاهه با نظر متخصصان این حوزه تأیید شده است. جامعه این پژوهش نظام‌های بازیابی اطلاعات شناخته‌شده دنیا هستند که از ابزارهای معنایی (شامل فرهنگ‌ها، سرعنوان‌های موضوعی، اصطلاح‌نامه‌ها و هستی‌شناسی‌ها) در بازیابی اطلاعات استفاده می‌کنند. چون این بررسی به‌صورت مقایسه‌ای انجام نگرفته و هدف مطالعه، نمایش ویژگی‌های معنایی متفاوت موجود در نظام‌های بازیابی اطلاعات مختلف، اعم از فهرست‌های پیوسته کتابخانه‌ها، پایگاه‌های اطلاعاتی و مراکز اسناد و آرشیوی و ... است، از هر یک از موارد موجود در جهان و ایران نمونه‌هایی انتخاب شد و مورد بررسی قرار گرفت.

جدول ۱. لیست نظام‌ها برای بررسی ویژگی‌های معنایی

ردیف	نام نظام بازیابی اطلاعات	ویژگی خاص برای انتخاب
۱	وبگاه کتابخانه کنگره آمریکا	استفاده از سرعنوان‌های موضوعی و اصطلاح‌نامه‌های مختلف در بازیابی اطلاعات
۲	وبگاه کتابخانه ملی استرالیا	امکان جست‌وجو بر اساس اصطلاح‌نامه
۳	وبگاه آرشیو ملی استرالیا	استفاده از مدل‌های مفهومی در بازیابی اطلاعات
۴	وبگاه موزه بریتانیا	استفاده از اصطلاح‌نامه مواد در سازماندهی و بازیابی اطلاعات
۵	پایگاه اطلاعاتی اریک	استفاده از اصطلاح‌نامه در جست‌وجو
۶	پایگاه مرکز تحقیقات کامپیوتری علوم اسلامی (نور)	استفاده از اصطلاح‌نامه در سازماندهی اطلاعات
۷	پایگاه پابان‌نامه‌های پژوهشگاه علوم و فناوری اطلاعات ایران (گنج)	استفاده از اصطلاح‌نامه در نمایه‌سازی

در میان نظام‌های اشاره‌شده در جدول ۱، «پایگاه گنج ایرانداک» از اصطلاح‌نامه تنها در نمایه‌سازی به‌عنوان مرحله‌ای از سازماندهی اطلاعات استفاده می‌کند و شواهد آن در بازیابی اطلاعات مشهود نیست، بنابراین، از لیست بررسی حذف گردید.

با مطالعه متون و منابع موجود در این زمینه مشخص شد که محققان مختلف ویژگی‌های مختلفی را برای یک جست‌وجوی معنایی مورد توجه قرار داده‌اند. این ویژگی‌ها شامل پنج مقوله کلی و مقوله‌های فرعی مربوط به آن‌ها هستند. در ویژگی‌های مربوط به جست‌وجوی مفهومی ۲۹ مقوله فرعی شناسایی شد. در گزینه‌های مربوط به پیشنهاد برای جست‌وجو ۳ مقوله، در گزینه‌های مربوط به جست‌وجوی مترادف‌ها ۲ مقوله و در گزینه‌های مربوط به نمایش ۹ مقوله ارائه و مورد بررسی قرار گرفت.

در گروه ویژگی‌های مربوط به هستی‌شناسی‌ها چهار گزینه برای بررسی بر اساس مطالعات تعریف شد. مقوله خاصی که در این پژوهش بر اساس مطالعه Bates (1990) مورد توجه قرار گرفت مربوط به تاکتیک‌هایی است که می‌تواند به‌طور مؤثر با هستی‌شناسی‌ها با حرکت درون هستی‌شناسی حمایت شود. این تاکتیک‌ها شامل ۱۰ تاکتیک هستند که می‌توانند نشان‌دهنده استفاده از هستی‌شناسی و ویژگی‌های هستی‌شناسانه در نظام بازیابی اطلاعات باشند.

تعدادی از معیارها با بررسی نظام قابل مشاهده بود، اما برای بررسی سایر معیارها با توجه به ماهیت آن‌ها، پرس‌و‌جو‌هایی بر اساس نمونه‌های موجود تعریف شد و در نظام بازیابی مورد نظر مشاهده گردید. نتایج این مشاهدات به‌صورت صفر و یک در جدولی مشخص گردید.

۵. نظام‌های بازیابی اطلاعات

کتابخانه کنگره آمریکا

فهرست آنلاین «کتابخانه کنگره آمریکا» یکی از شناخته‌شده‌ترین نظام‌های بازیابی اطلاعات جهان است که در بازیابی اطلاعات خود از اصطلاح‌نامه‌ها و سرعنوان‌های موضوعی استفاده می‌کند. فهرست‌نویسی موضوعی به لیستی از موضوعات کنترل‌شده در دسترس اصطلاح‌نامه‌ها و واژگان وابسته است. رکوردهای فهرست «کتابخانه کنگره» عموماً نقاط دسترسی موضوعی را از یک یا چند اصطلاح‌نامه و سرعنوان‌های موضوعی زیر می‌گیرد.

- ◇ سرعنوان‌های موضوعی «کتابخانه کنگره»؛
- ◇ اصطلاحات سبک هنری/فرم^۱ «کتابخانه کنگره» برای مواد آرشیوی و کتابخانه‌ای؛
- ◇ اصطلاح‌نامه متوسط عملکرد موسیقی «کتابخانه کنگره»؛
- ◇ اصطلاح‌نامه‌ای برای مواد گرافیکی: اصطلاحات موضوعی و ژانری؛
- ◇ واژگان کنترل‌شده: RBMS^۲ واژگان کنترل‌شده برای استفاده در کتب نایاب و فهرست‌نویسی مجموعه‌های خاص؛
- ◇ راهنماهایی در دسترسی موضوعی به آثار شخصی، داستان، درام و ...؛
- ◇ راهنمای ژانر/ فرم تصاویر متحرک، راهنمای اصطلاحات سبک هنری/ فرم رادیو، اصطلاحات موضوعی قانون‌گذاری (Thesauri & Controlled Vocabularies 2016).

در سیستم جست‌وجوی کاتالوگ‌های «کتابخانه کنگره» از این ساختارهای اصطلاح‌نامه‌ای و همچنین از سرعنوان‌های موضوعی برای بهبود نتایج جست‌وجو بهره گرفته می‌شود. به‌عنوان مثال، اگر در جست‌وجوی پیشرفته عبارت Iranian Revolution و Revolution را به‌عنوان موضوع جست‌وجو وارد کنیم، این درخواست به مجموعه سرعنوان موضوعی خواهد رفت و عبارت مناسب با آن با استفاده از ساختار درختی موجود استخراج می‌شود و به عبارات درخواستی کاربر افزوده می‌شود و جست‌وجو با همه این عبارات انجام خواهد شد. برای مثال، یکی از نتایج کتابی است از خاطرات جبهه مربوط به «لشکر امام حسین اصفهان» که در تقسیمات فرعی مجموعه سرعنوان‌ها قرار گرفته است. پس از جست‌وجو در اطلاعات کتابشناختی، در قسمت موضوعی که به مدرک تعلق گرفته، محل قرار گرفتن موضوع مدرک را در مجموعه سرعنوان‌های موضوعی مشخص می‌کند که می‌توان با کلیک کردن روی آن، مجموعه رکوردهای جدیدی را با اصطلاح جست‌وجوی جدید در یک لیست مشاهده کرد (استفاده از اصطلاح‌نامه در پایگاه‌های اطلاعاتی مرکز اطلاعات و مدارک علمی ایران ۱۳۸۳).

کتابخانه ملی استرالیا

بزرگ‌ترین کتابخانه و یکی از مهم‌ترین مراکز پژوهشی استرالیا، کتابخانه ملی آن است. سابقه فعالیت این کتابخانه به سال ۱۹۰۲ باز می‌گردد. وب‌گاه «کتابخانه استرالیا»

1. Genre/ Form

2. Rare Book and Manuscript section

شامل جست‌وجوی ساده، پیشرفته و مرور در فهرست «کتابخانه ملی استرالیا» و سایر کتابخانه‌های استرالیا از جمله کتابخانه‌های عمومی است. همچنین، می‌توان در منابع دیجیتالی کتابخانه و نیز در منابع الکترونیکی آن جست‌وجو انجام داده و به اطلاعات مورد نیاز دست یافت.

«کتابخانه ملی استرالیا» از اصطلاح‌نامه APAIS استفاده می‌کند. این اصطلاح‌نامه اصطلاحات موضوعی مورد استفاده برای نمایه‌سازی مقالات برای APAIS را لیست می‌کند. این اصطلاح‌نامه در سال ۱۹۷۸ وقتی که APAIS به‌عنوان پایگاه اطلاعاتی آنلاین در دسترس قرار گرفت، آغاز به کار کرد و اولین بار در سال ۱۹۸۰ توسط «کتابخانه ملی استرالیا» برای تسهیل جست‌وجو ایجاد شد.

این اصطلاح‌نامه هنوز به‌عنوان ابزار نمایه‌سازی استفاده شده و ابزار کمکی برای جست‌وجوی آنلاین APAIS و پایگاه اطلاعاتی تمام متن روابط عمومی استرالیا است. علاوه بر خدمت به مشتریان APAIS اصطلاح‌نامه منتشر شده در تعدادی پروژه که نیاز به فهرست ساختار بندی شده اصطلاحات موضوعی استرالیا دارند، نیز مورد استفاده قرار گرفت (National library of Australia 2016).

آرشیو ملی استرالیا

«آرشیو ملی استرالیا» از اصطلاح‌نامه عملکرد تعاملی دولت استرالیا (AGIFT) در بازیابی اطلاعات بهره می‌گیرد. این اصطلاح‌نامه چارچوبی استاندارد برای توصیف عملکردهای دولت (AGIFT) است. این اصطلاح‌نامه، اصطلاح‌نامه‌ای با سلسله‌مراتب سه سطحی است که عملکردهای تجاری انجام شده در نواحی مشترک‌المنافع، ایالات و دولت‌های محلی در استرالیا را توصیف می‌کند. اهداف این اصطلاح‌نامه شامل موارد زیر است:

- ◇ فراهم‌آوری اصطلاحات استاندارد برای ادارات دولتی برای استفاده در عنصر «عملکرد» مجموعه عناصر فراداده‌ای^۱ (AGLS)؛
- ◇ کمک به کاربران در جست‌وجوی مدخل‌هایی که مطمئن نیستند کدام اصطلاح مورد استفاده قرار گرفته یا کدام سطح دولتی برای اطلاعات یا خدمات مورد نیاز مسئول است؛

1. Australian Government Locator Service

◇ فراهم‌آوری چارچوبی برای ادارات دولتی برای توسعه اصطلاح‌نامه‌های عملکردی مبتنی بر اداره برای نیازهای طبقه‌بندی آن‌ها.

کتابخانه موزه بریتانیا

مجموعه پایگاه اطلاعاتی از یک سیستم از اصطلاحات کنترل‌شده برای نمایه‌سازی و طبقه‌بندی مجموعه موزه بریتانیا استفاده می‌کند. این‌ها شامل انواع اشیاء مثل اشیای چاپی، مجسمه‌سازی یا کاسه و مکان‌های مرتبط به این اشیاء می‌شود. این اصطلاحات کنترل‌شده در اصطلاح‌نامه و فایل‌های مستند آمده‌اند که به موزه‌داران و سایر کاربران پایگاه اطلاعاتی اجازه دسته‌بندی یکدست اشیاء را بدهند.

یکی از اصطلاح‌نامه‌ها، اصطلاح‌نامه‌ای برای نام‌های اشیاء و مواد است. این‌ها با اعتماد به مجموعه‌ها منتشر شده (مانند MDA/انجمن دکومانتاسیون موزه) و با دیگر موزه‌ها در سراسر جهان پذیرفته شده است.

پایگاه اطلاعاتی اریک

«مرکز اطلاعاتی منابع تعلیم و تربیت» (اریک) دسترسی نامحدود به بیش از ۱/۲ میلیون سند کتابشناختی از مقالات نشریات و سایر مطالب مربوط به تعلیم و تربیت را فراهم می‌کند و هر دو هفته یک‌بار صدها سند جدید به آن اضافه می‌شود. در صورت دسترسی، لینک‌ها به متن کامل مقالات هم اضافه می‌شود.

«اریک» مرکز منابع اطلاعات علمی و آموزش و پرورش است و نیز یک کتابخانه دیجیتالی آنلاین با اطلاعات گسترده علمی و پژوهشی. این پایگاه امکان دستیابی به اطلاعات را جهت پشتیبانی و استفاده از پژوهش‌ها و نیز در آموزش، تدریس و پژوهش فراهم می‌کند.

در این پایگاه علاوه بر امکان جست‌وجوی ساده و پیشرفته، امکان جست‌وجو در اصطلاح‌نامه این پایگاه نیز وجود دارد. کاربرانی که مهارت استفاده از جست‌وجوی پیشرفته را ندارند یا می‌خواهند مناسب‌ترین عبارت را برای جست‌وجوی مورد نظر خود بیابند یا می‌خواهند با یافتن عبارات مرتبط با موضوع، حجم اطلاعات بازیابی شده را افزایش دهند، می‌توانند از طریق جست‌وجو در اصطلاح‌نامه، موضوعات مورد علاقه خود را یافته و اطلاعاتی در مورد آن بازیابی کنند. بنابراین، بسط درخواست کاربر و افزودن عبارات هم‌معنا و متناسب به صورت اتوماتیک متفی است و این وظیفه به خود کاربر

واگذار شده است (استفاده از اصطلاحنامه در پایگاه‌های اطلاعاتی مرکز اطلاعات و مدارک علمی ایران ۱۳۸۳).

پایگاه اطلاعاتی مجلات نورمگز

«مرکز تحقیقات کامپیوتری علوم اسلامی»، پیشگام در رایانه‌ای کردن منابع دینی و پژوهش‌های اسلامی است. داده‌پردازی علوم اسلامی و به کارگیری رایانه در پژوهش‌های دینی و اسلامی از همان ابتدا در رأس فعالیت‌های این مرکز بوده است. در حال حاضر، این مرکز در همهٔ موضوعات علوم اسلامی نرم‌افزارهایی را با رویکردهای پژوهشی، تبلیغی و اطلاع‌رسانی برای سطوح تخصصی و عمومی تولید کرده است. هر یک از پایگاه‌های مرکز برای اهدافی خاص یا در شرایطی معین طراحی و تولید شده‌اند. «پایگاه اطلاعاتی مجلات نورمگز» نیز جزو این پایگاه‌هاست. «شبکهٔ جهانی نور» از سال ۱۳۷۷ اقدام به عرضهٔ مجلات مربوط به علوم اسلامی و علوم انسانی بر روی پایگاه حوزه نمود. این پایگاه از مهرماه سال ۱۳۸۴ کار خود را آغاز کرده که ایجاد بایگانی مجلات و مقالات علوم انسانی و اسلامی در اینترنت یکی از مهم‌ترین اهداف راه‌اندازی پایگاه مجلات تخصصی بوده است. درج بیش از ۳۲۰۰۰ شماره از ۷۰۰ عنوان مجلهٔ تخصصی در قالب موضوعات مختلفی چون «قرآن و حدیث»، «فلسفه، کلام و عرفان»، «فقه، اصول و حقوق»، «کناب‌داری و اطلاع‌رسانی»، «روان‌شناسی»، «اقتصاد»، «هنر»، «مدیریت»، «حقوق»، «علوم انسانی» و «تاریخ»، این پایگاه را به یکی از مهم‌ترین بانک‌های اطلاعات مجلات در زمینهٔ علوم اسلامی و انسانی تبدیل کرده است (اکبری ۱۳۸۹).

۶. تجزیه و تحلیل یافته‌های پژوهش

جست‌وجوی مفهومی

جست‌وجوی مفهومی با جست‌وجوی فراداده‌ای یا کلیدواژه‌ای تفاوت بسیاری دارد. جست‌وجوی مفهومی در تحلیل نسبی، مطابقت مفهومی بین پرس‌وجو و سند را نشان می‌دهد. در واقع، این نوع جست‌وجو، فنی است که کلماتی را که با مفهوم یک کلمهٔ پرس‌وجو مشابه هستند، فراهم می‌کند. این نوع جست‌وجو اسنادی را بازیابی می‌کند که با همان مفهوم کلمهٔ جست‌وجو مرتبط است، بدون توجه به این که کلمهٔ پرس‌وجو در اسناد نتایج جست‌وجو وجود داشته باشد یا نه.

در این بخش و با توجه به تعاریف و ویژگی‌هایی که متون برای این نوع جست‌وجو ارائه کرده‌اند، ۲۹ گزینه مربوط به ویژگی‌های آن شناسایی شد و در نظام‌های بازیابی اطلاعات مورد بررسی قرار گرفت.

جدول ۲. معیارهای جست‌وجوی مفهومی در نظام‌های بازیابی اطلاعات

ردیف	معیارهای جست‌وجوی مفهومی	کتابخانه کنگره آمریکا	کتابخانه ملی استرالیا	آرشیو ملی استرالیا	موزه بریتانیا	پایگاه اطلاعاتی اریک	پایگاه مرکز تحقیقات کامپیوتری علوم اسلامی (نور)
۱	استفاده از واژگان کنترل‌شده	۱	۱	۱	۱	۱	۱
۲	توسعه اصطلاح‌نامه‌ای	۰	۰	۰	۱	۱	۱
۳	جست‌وجوی هستی‌شناسی	۰	۰	۰	۰	۰	۰
۴	استفاده دستی از اصطلاح‌نامه	۱	۰	۰	۰	۱	۰
۵	جست‌وجوی مترادف‌ها با اصطلاح‌نامه	۱	۰	۰	۰	۱	۰
۶	ارائه مجموعه‌ای از مفاهیم در حوزه‌ای خاص	۰	۱	۱	۱	۱	۰
۷	انتخاب دامنه	۰	۰	۱	۰	۱	۱
۸	انتخاب مفاهیم سطح بالا در هر دامنه	۰	۰	۱	۰	۱	۰
۹	تنوع ریخت‌شناختی نتایج در واژه جست‌وجوشده مانند حالت‌های مختلف زمان فعل، و جمع و مفرد بودن	۱	۱	۰	۰	۱	۱
۱۰	امکان عمومیت‌بخشی	۰	۱	۰	۱	۱	۰
۱۱	توجه به تغییرات مورفولوژیکی	۱	۱	۰	۰	۰	۰
۱۲	در نظر گرفتن آیتم‌های شناخته‌شده	۱	۰	۰	۰	۰	۰
۱۳	ارائه ارتباطات میان اصطلاحات	۰	۰	۰	۰	۱	۰

ردیف	معیارهای جست و جوی مفهومی	کتابخانه کنگره آمریکا	کتابخانه ملی استرالیا	آرشیو ملی استرالیا	موزه بریتانیا	پایگاه اطلاعاتی اریک	پایگاه مرکز تحقیقات کامپیوتری علوم اسلامی (نور)
۱۴	ارائه نتایج مرتبط با نیاز اطلاعاتی بیان شده	۱	۱	۱	۱	۱	۱
۱۵	رتبه‌بندی و ذخیره بر اساس ارتباط	۱	۱	۱	۱	۱	۱
۱۶	تعیین محل و نمایش سریع نتایج	۱	۱	۱	۱	۱	۱
۱۷	ثابت نبودن طول پرس و جو	۱	۰	۰	۰	۱	۰
۱۸	امکان مرور سلسله‌مراتبی	۱	۱	۱	۰	۱	۰
۱۹	ایجاد سریع نمایه‌های پرس و جوی آماده	۰	۰	۰	۰	۱	۱
۲۰	عدم نیاز به نحو پیچیده و خاص	۱	۱	۱	۱	۱	۱
۲۱	اجازه پرس و جوی ترکیبی با استفاده از مفاهیم، کلیدواژه‌ها، و فراداده‌ها	۱	۱	۱	۱	۱	۰
۲۲	عدم تأثیرپذیری از املا، نادرست کلمات، اشتباهات تایپی یا اشتباهات اسکن OCR در متن پرس و جو	۰	۰	۰	۰	۱	۰
۲۳	بازایی نتایج مرتبط بدون توجه به وجود یا نبود مفهوم در پرس و جو	۱	۰	۰	۰	۱	۰
۲۴	استفاده از ریشه لغت	۰	۰	۱	۰	۱	۱
۲۵	استخراج معنای بالقوه پرس و جو	۱	۰	۰	۰	۱	۰
۲۶	همخوانی با تغییرات معنایی اصطلاحات	۱	۱	۰	۰	۰	۰
۲۷	مشخص کردن تعداد نتایج	۱	۰	۰	۰	۰	۰
۲۸	استخراج موجودیت، نام‌های خاص و دیگر اطلاعات خاص برای اهداف جست و جو	۰	۱	۰	۰	۰	۰

ردیف	معیارهای جست‌وجوی مفهومی	کتابخانه کنگره آمریکا	کتابخانه ملی استرالیا	آرشیو ملی استرالیا	موزه بریتانیا	پایگاه اطلاعاتی اریک	پایگاه مرکز تحقیقات کامپیوتری علوم اسلامی (نور)
۲۹	امکان انتخاب معنای مورد نظر در مورد چندمعنایی‌ها	۰	۰	۰	۰	۰	۰
	کل	۱۷	۱۳	۱۱	۹	۲۲	۱۰
	درصد	۵۸/۶۲	۴۴/۸۲	۳۷/۹۳	۳۱/۰۳	۷۵/۸۶	۳۴/۴۸
	میانگین	۱۳/۶۶					

بررسی ۲۹ ویژگی مربوط به جست‌وجوی مفهومی، همان‌طور که در جدول ۲، مشخص است، نشان می‌دهد که «پایگاه اطلاعاتی اریک» با داشتن ۲۲ امتیاز و ۷۵/۸۶ درصد از امتیاز معیارهای مشخص شده بیشتر از سایر نظام‌ها از جست‌وجوی مفهومی پشتیبانی می‌کند. پس از «اریک»، «کتابخانه کنگره آمریکا» با ۱۷ امتیاز و ۵۸/۶۲ درصد از امتیاز معیارهای جست‌وجوی مفهومی در رتبه دوم قرار دارد. «موزه بریتانیا» با ۹ امتیاز و ۳۱/۰۳ درصد از این معیارها در رتبه آخر قرار گرفته است. امتیاز میانگین کل نظام‌های بازیابی اطلاعات ۱۳/۶۶ است.

همان‌طور که در جدول مشخص است، ویژگی‌های استفاده از واژگان کنترل‌شده، اجازه پرس‌وجوی ترکیبی با استفاده از مفاهیم، کلیدواژه‌ها و فراداده‌ها، ارائه نتایج مرتبط با نیاز اطلاعاتی بیان‌شده، رتبه‌بندی و ذخیره بر اساس ارتباط و تعیین محل و نمایش سریع نتایج در تمامی نظام‌ها به صورت مشترک وجود داشته است. ارائه مجموعه‌ای از مفاهیم در حوزه‌ای خاص و توسعه اصطلاح‌نامه‌ای نیز در میان ویژگی‌های جست‌وجوی مفهومی در بیشتر نظام‌ها (۴ نظام از ۶ نظام مورد نظر) به کار رفته است.

پیشنهاد برای جست‌وجو

زمانی که جست‌وجوگر جست‌وجوی اولیه خود را انجام می‌دهد، از وسعت واژگانی که می‌تواند برای جست‌وجوی خود استفاده کند، آگاه نیست و اغلب جست‌وجو را به کلیدواژه‌های خاص و در زبان طبیعی محدود می‌کند. پیشنهاد برای جست‌وجو به کاربر اجازه می‌دهد که مفاهیم مرتبط با کلیدواژه‌های مورد جست‌وجوی خود را ببیند و در صورت لزوم جست‌وجوی خود را بر اساس آن انجام دهد.

در این معیار، ۳ مقوله مد نظر قرار گرفت و در نظام‌های بازیابی اطلاعات مورد نظر بررسی شد.

جدول ۳. معیارهای پیشنهاد برای جست‌وجو در نظام‌های بازیابی اطلاعات

معیارهای پیشنهاد برای جست‌وجو	کتابخانه کنگره آمریکا	کتابخانه ملی استرالیا	آرشیو ملی استرالیا	موزه بریتانیا	پایگاه اطلاعاتی اریک	پایگاه مرکز تحقیقات کامپیوتری علوم اسلامی (نور)
ارائه پیشنهاد بر اساس اصطلاحات نتایج مرتبط	۰	۱	۰	۰	۰	۰
ارائه پیشنهاد بر اساس اصطلاح‌نامه	۰	۰	۰	۰	۰	۱
پیشنهاد رفع ابهام اصطلاحات پرس‌وجو	۰	۰	۰	۰	۰	۰
کل	۰	۱	۰	۰	۰	۱
درصد میانگین	۰/۳۳	۳۳/۳۳	۰	۰	۰	۳۳/۳۳

همان‌طور که در جدول ۳، مشخص است، اغلب نظام‌های مورد بررسی در این پژوهش معیارهای پیشنهاد برای جست‌وجو ندارند و در به کار بردن این ویژگی‌ها ضعیف هستند. «کتابخانه ملی استرالیا» نتایج مشابه با نتایج بازیابی شده را در فهرستی ارائه می‌دهد.

جست‌وجوی مترادف‌ها

جست‌وجوی مترادف‌ها با استفاده از اصطلاح‌نامه اسنادی را که دارای اصطلاحات مشابه در معنا با کلمات پرس‌وجو هستند، بازیابی می‌کند. در این معیار، ۲ مقوله مد نظر قرار گرفت و در نظام‌های بازیابی اطلاعات مورد نظر بررسی شد.

جدول ۴. معیارهای جست‌وجوی مترادف‌ها در نظام‌های بازیابی اطلاعات

معیارهای جست‌وجوی مترادف‌ها	کتابخانه کنگره آمریکا	کتابخانه ملی استرالیا	آرشیو ملی استرالیا	موزه بریتانیا	پایگاه اطلاعاتی اریک	پایگاه مرکز تحقیقات کامپیوتری علوم اسلامی (نور)
جست‌وجوی مترادف‌ها با اصطلاح‌نامه	۱	۰	۰	۰	۱	۱
جست‌وجوی مترادف‌ها با هستی‌شناسی	۰	۰	۰	۰	۰	۰
کل	۱	۰	۰	۰	۱	۱
درصد	۵۰	۰	۰	۰	۵۰	۵۰
میانگین	۰/۵					

به این دلیل که اغلب نظام‌ها از هستی‌شناسی در بازیابی اطلاعات خود استفاده نمی‌کردند، معیار جست‌وجوی مترادف‌ها بر اساس اصطلاح‌نامه مد نظر قرار می‌گیرد و همان‌طور که در جدول ۴، آمده، تنها سه مورد از نظام‌ها این امر را مورد توجه قرار داده بودند. امتیاز میانگین ۰/۵ از ۶ امتیاز کل، امتیاز بسیار پایینی است. ویژگی استفاده از اصطلاح‌نامه‌ها در جست‌وجوی مترادف‌ها در نیمی از نظام‌ها به کار رفته است و حضور آن در پرداختن به ویژگی‌های معنایی اهمیت زیادی دارد.

استفاده از هستی‌شناسی

در نظام‌های بازیابی اطلاعات استفاده از هستی‌شناسی به‌عنوان راه حلی برای مشکلات جست‌وجوی کلیدواژه‌ای و تبدیل جست‌وجوی سنتی به جست‌وجوی معنایی پیشنهاد شده است. بنابراین، استفاده از آن می‌تواند در بازیابی معنایی مفید باشد. به‌طور کلی، چهار مقوله کلی برای این معیار شناسایی شد. معیار استفاده از هستی‌شناسی در تاکتیک‌های جست‌وجو شامل ۱۰ مقوله فرعی بوده است.

«بیتس» (۱۹۹۰) اشاره می‌کند که گاهی کاربر ترجیح می‌دهد که فرایند جست‌وجو را تا حدی بر طبق استراتژی‌ها و تاکتیک‌های مشخصی که بتواند به‌طور مؤثر در رابط کاربر حمایت شود، کنترل کند. او تاکتیک‌ها را به‌عنوان «یک یا تعداد انگشت‌شماری از حرکاتی که جست‌وجو را پیش می‌برند» تعریف می‌کند و به‌عنوان مثال، آن‌ها را شامل

تغییر یک اصطلاح به اصطلاحات خاص تر یا اضافه کردن اصطلاحات مشابه می‌داند (Bates 1990). در واقع، هستی‌شناسی‌ها می‌توانند برای حمایت مستقیم تاکتیک‌ها استفاده شوند (García and Sicilia 2003).

جدول ۵، خلاصه‌ای از تاکتیک‌های اشاره‌شده در پژوهش (Bates 1990) را که می‌تواند به‌طور مؤثر از طریق هستی‌شناسی‌ها حمایت شود، توصیف می‌کند. در این جدول حروف بزرگ (A,B,C) برای مشخص کردن اصطلاحات است (García and Sicilia 2003).

جدول ۵. تاکتیک‌های جست‌وجوی حمایت‌شونده از سوی هستی‌شناسی (García and Sicilia 2003)

تاکتیک	توصیف بی‌س	پیاده‌سازی مبتنی بر هستی‌شناسی	مثال
SUPER	حرکت سلسله‌مراتبی به اصطلاح اعم (بالا تر)	تغییر B با A، B زیرمجموعه A	حرکت از گره چشم‌آبی به گره
SUB	حرکت سلسله‌مراتبی به اصطلاح اخص (فرعی)	تغییر A با B، B زیرمجموعه A است.	حرکت از گره به گره چشم‌آبی
RELATE	حرکت یک‌طرفه سلسله‌مراتبی به هم‌پایه	تغییر B با C جایی که مفهوم A وجود دارد؛ به طوری که B زیرمجموعه A، C زیرمجموعه A	حرکت از گره چشم‌آبی به گره براق از طریق گره
CONTRARY	جست‌وجوی اصطلاحی که صورت منطقی مخالف توصیف مورد نظر است	تغییر B با C جایی که اشتراک B و C تهی است (حداقل)	حرکت از محصول نهایی به محصول اولیه (به عنوان دسته‌بندی مشروبات)
RECORD	پیگیری مسیرهای آزمایشی که دنبال شده و آزمایش‌های مطلوبی که دنبال یا تکمیل نشده	ذخیره هستی‌شناسی مسیرهای مشاهده‌شده در حال مرور	مسیر حیوان خانگی به عنوان یک اصطلاح به گره (اصطلاح بعدی) و سپس به گره چشم‌آبی ثبت شده. بنابراین، سیستم ممکن است بعداً به دیگر مشاهده‌ها پیشنهاد دهد
SELECT	تجزیه پرس‌وجوی جست‌وجوی پیچیده به مشکلات فرعی و کاربر بر روی یک مشکل در یک زمان	تجزیه جست‌وجوی مداوم، ویرایش بعضی از شاخه‌هایی که می‌توانند بعداً اضافه شوند	وقتی که جست‌وجو برای اطلاعاتی در مورد ابزارهای آزمایش نرم‌افزار انجام می‌شود، تجزیه فنون مختلف (واحد، یکپارچه و ...)

تاکتیک	توصیف پیش	پایاده‌سازی مبتنی بر هستی‌شناسی	مثال
PARALLEL	گسترش فرمول‌بندی جست‌وجو (یا گسترده‌تر) با شامل کردن مترادف‌ها یا دیگر اصطلاحات برابر مفهومی	افزودن اصطلاح C وقتی که C مشابه با هر اصطلاحی که بخشی از جهت‌یابی جاری را تشکیل می‌دهد	وقتی جست‌وجو برای اطلاعات در مورد بازرسی نرم‌افزار شامل نقد نرم‌افزار نیز می‌شود.
PINPOINT	ایجاد فرمول‌بندی جست‌وجوی دقیق با حداقل سازی (یا کاهش) تعداد اصطلاحات برابر، حفظ اصطلاحات توصیفی‌تر	مخالف PARALLEL -	
SPECIFY	جست‌وجوی اصطلاحاتی که به اندازه اطلاعات مطلوب خاص هستند	اجازه اصلاح قدم‌به‌قدم با پیش بردن به سوی خاص تر شدن در سلسله‌مراتب تابعی، ترکیب چند اصطلاح با «و» شبیه معناها	بعداً به‌عنوان اصول طراحی راهنمای اصلی ONTOIR توضیح داده خواهد شد.
EXHAUST	شامل بیشتر یا تمام عناصر پرس‌وجو در فرمول‌بندی جست‌وجوی اولیه: برای افزودن یک یا چند عنصر پرس‌وجو به فرمول جست‌وجوی آماده	اجازه به انتخاب چند اصطلاح و «یا»- شبیه معناها	بعداً به‌عنوان ویژگی ONTOIR توضیح داده خواهد شد.

تاکتیک‌های جست‌وجوی مبتنی بر هستی‌شناسی طبق جدول ۵، مورد بررسی قرار گرفت. نتایج بررسی این مقوله‌ها در نظام‌ها در جدول ۶، آمده است.

جدول ۶. معیارهای استفاده از هستی‌شناسی در نظام‌های بازایی اطلاعات

معیارهای استفاده از هستی‌شناسی	کتابخانه کنگره آمریکا	کتابخانه ملی استرالیا	آرشیو ملی استرالیا	موزه بریتانیا	پایگاه اطلاعاتی اریک	پایگاه مرکز تحقیقات کامپیوتری علوم اسلامی (نور)
استفاده از رابط کاربر برای تایپ اصطلاحات در هستی‌شناسی	•	•	•	•	•	•
استفاده از هستی‌شناسی برای توسعه پرس‌وجو	•	•	•	•	•	•
ارائه ساختار فنی هستی‌شناسی در رابط کاربر	•	•	•	•	•	•

معیارهای استفاده از هستی‌شناسی	کتابخانه کنگره آمریکا	کتابخانه ملی استرالیا	آرشیو ملی استرالیا	موزه بریتانیا	پایگاه اطلاعاتی اریک	پایگاه مرکز تحقیقات کامپیوتری علوم اسلامی (نور)
SUPER	۰	۱	۰	۱	۱	۰
SUB	۰	۱	۰	۱	۱	۰
RELATE	۰	۱	۰	۰	۱	۰
CONTRARY	۰	۰	۰	۰	۱	۰
RECORD	۰	۰	۰	۰	۰	۰
SELECT	۰	۰	۰	۰	۰	۰
PARALLEL	۰	۰	۰	۰	۰	۰
PINPOINT	۰	۰	۰	۰	۰	۰
SPECIFY	۰	۱	۰	۰	۰	۰
EXHAUST	۰	۱	۰	۰	۱	۰
کل	۰	۶	۰	۳	۶	۰
درصد	۰	۴۲/۸۵	۰	۲۱/۴۲	۴۲/۸۵	۰
میانگین	۱/۸۷۵					

همان‌طور که در جدول ۶، آمده، نظام‌های بازیابی اطلاعات ویژگی‌های هستی‌شناسانه را در بازیابی خود خیلی کم به کار برده‌اند. «کتابخانه ملی استرالیا» و «پایگاه اطلاعاتی اریک» با ۴۲/۸۵ درصد رتبه اول را دارند، اما این میزان استفاده کمتر از نیمی از ویژگی‌هاست. در میان تاکتیک‌های جست‌وجوی مبتنی بر هستی‌شناسی، تاکتیک SUPER و SUB (در ۳ نظام از ۶ نظام) که به‌نوعی اصطلاحات اعم و اخص را نشان می‌دهد، بیشتر از سایر تاکتیک‌ها در نظام‌ها به کار گرفته شده است و این نشان‌دهنده محدود بودن این نظام‌ها به روابط اصطلاح‌نامه‌ای است.

نمایش

استفاده از ویژگی‌های معنایی در بازیابی اطلاعات، نمودهایی در نمایش دارد. این معیارهای استخراج‌شده از متون شامل ۹ معیار است. این معیارها در هر یک از نظام‌های بازیابی اطلاعات بررسی شد.

جدول ۷. معیارهای نمایش در نظام‌های بازایی اطلاعات

معیارهای نمایش	کتابخانه کنگره آمریکا	کتابخانه ملی استرالیا	آرشیو ملی استرالیا	موزه بریتانیا	پایگاه اطلاعاتی اریک	پایگاه مرکز تحقیقات کامپیوتری علوم اسلامی (نور)
ارائه بخش مرتبط به نتایج به شکل ساده	۱	۱	۱	۱	۱	۱
نمایش موضوعات مرتبط	۰	۱	۰	۰	۰	۱
نمایش برجسته کلیدواژه‌های مرتبط به با پرس و جو	۱	۰	۰	۰	۱	۱
تعمیم یک مفهوم کلی به اجزای آن	۰	۱	۰	۰	۰	۰
ارائه مدلی تصویری از نتایج کاوش	۰	۰	۰	۰	۰	۰
نمایش رکورد به همراه خلاصه	۱	۱	۱	۱	۱	۱
ارائه نتایج یکسان برای پرس و جوهای حاوی نحو متفاوت و معنای مشابه	۰	۰	۰	۰	۰	۰
نمایش اشکال مختلف دستوری پرس و جو	۰	۰	۰	۰	۰	۰
ارائه تعریف	۰	۰	۰	۰	۰	۰
کل	۲	۴	۲	۲	۳	۴
درصد	۲۲/۲۲	۴۴/۴۴	۲۲/۲۲	۲۲/۲۲	۳۳/۳۳	۴۴/۴۴
میانگین	۱/۳					

همان‌طور که در جدول ۷، مشخص است، نظام‌های بازایی اطلاعات مورد بررسی در این پژوهش به‌ندرت در نمایش خود ویژگی‌های معنایی را نشان دادند. در میان آن‌ها، «پایگاه مرکز تحقیقات کامپیوتری علوم اسلامی» و «کتابخانه ملی استرالیا» تعداد بیشتری از این معیارها (۴ معیار از ۹ معیار، ۴۴/۴۴ درصد) را دارا بودند. در این میان «آرشیو ملی استرالیا» و «موزه بریتانیا» از این معیارها خیلی کم (۲ معیار از ۹ معیار، ۲۲/۲۲ درصد) استفاده کردند.

ارائه بخش مرتبط با نتایج به شکل ساده و نمایش رکورد به همراه خلاصه در تمامی نظام‌ها وجود داشت. نمایش برجسته کلیدواژه‌های مرتبط به با پرس‌وجو که نمونه‌ای ساده از نمایش ویژگی‌های معنایی است، تنها در نیمی از نظام‌ها مورد توجه قرار گرفته است.

جدول ۸. ویژگی‌های معنایی کلی در نظام‌های بازیابی اطلاعات

معیارهای معنایی	کتابخانه کنگره آمریکا	کتابخانه ملی استرالیا	آرشیو ملی استرالیا	موزه بریتانیا	پایگاه اطلاعاتی اریک	پایگاه مرکز تحقیقات کامپیوتری علوم اسلامی (نور)
امتیاز کلی	۲۰	۲۴	۱۳	۱۴	۳۲	۱۶



نمودار ۱. ویژگی‌های معنایی در نظام‌های بازیابی اطلاعات

جدول ۸، و نمودار ۱، وضعیت نظام‌های مورد بررسی را از لحاظ ویژگی‌های معنایی و هستی‌شناسانه نشان می‌دهد. همان‌طور که مشخص است، «پایگاه اطلاعاتی اریک» و «کتابخانه ملی استرالیا» با ۳۲ و ۲۴ امتیاز از نظر ویژگی‌های معنایی در صدر نظام‌های بررسی شده قرار دارند و «آرشیو ملی استرالیا» و «موزه بریتانیا» نیز از این لحاظ با ۱۳ و ۱۴ امتیاز دارای کمترین امتیاز هستند.

۷. نتیجه‌گیری

در این پژوهش ویژگی‌های معنایی نظام‌های بازیابی مطرح دنیا که از اصطلاح‌نامه

و هستی‌شناسی در بازیابی اطلاعات خود استفاده می‌کنند، مورد بررسی قرار گرفت. هدف از این پژوهش ارزیابی این نظام‌ها نبود، بلکه تلاش شد ویژگی‌هایی که بیشتر در نظام‌ها به کار رفته است و ویژگی‌هایی که متون برای این نظام‌ها در نظر گرفته‌اند، شناسایی شود. ویژگی‌هایی که به‌طور کلی از متون به‌عنوان ویژگی‌های معنایی شناسایی شده، شامل این موارد است: جست‌وجوی مفهومی که خود شامل استفاده از واژگان کنترل‌شده (اصطلاح‌نامه، هستی‌شناسی‌ها، وردنت)، توسعه اصطلاح‌نامه‌ای، جست‌وجوی هستی‌شناسی، استفاده دستی از اصطلاح‌نامه، جست‌وجوی مترادف‌ها با اصطلاح‌نامه، ارائه مجموعه‌ای از مفاهیم در حوزه‌ای خاص، انتخاب دامنه، انتخاب مفاهیم سطح بالا در هر دامنه، تنوع ریخت‌شناختی نتایج در واژه جست‌وجوشده مانند حالت‌های مختلف زمان فعل، جمع و مفرد بودن، امکان عمومیت‌بخشی، توجه به تغییرات مورفولوژیکی، در نظر گرفتن آیم‌های شناخته‌شده، ارائه ارتباطات میان اصطلاحات، ارائه نتایج مرتبط با نیاز اطلاعاتی بیان‌شده، رتبه‌بندی و ذخیره بر اساس ارتباطات، تعیین محل و نمایش سریع نتایج، ثابت نبودن طول پرس‌وجو، امکان مرور سلسله‌مراتبی، ایجاد سریع نمایه‌های پرس‌وجوی آماده، عدم نیاز به نحو پیچیده و خاص، اجازه پرس‌وجوی ترکیبی با استفاده از مفاهیم کلیدواژه‌ها و فراداده‌ها، عدم تأثیرپذیری از املاهای نادرست کلمات، اشتباهات تایپی یا اشتباهات اسکن OCR در متن پرس‌وجو، بازیابی نتایج مرتبط بدون توجه به وجود یا نبود مفهوم در پرس‌وجو، استفاده از ریشه لغت، استخراج معنای بالقوه پرس‌وجو، همخوانی با تغییرات معنایی اصطلاحات، مشخص کردن تعداد نتایج، استخراج موجودیت، نام‌های خاص و دیگر اطلاعات خاص برای اهداف جست‌وجو، امکان انتخاب معنای مورد نظر در مورد چند معنایی‌ها. نتایج این بررسی نشان داد که بزرگ‌ترین نظام‌های بازیابی اطلاعات دنیا دارای ویژگی‌های معنایی شامل ویژگی‌های استفاده از واژگان کنترل‌شده، اجازه پرس‌وجوی ترکیبی با استفاده از مفاهیم، کلیدواژه‌ها، و فراداده‌ها، ارائه نتایج مرتبط با نیاز اطلاعاتی بیان‌شده، رتبه‌بندی و ذخیره بر اساس ارتباطات و تعیین محل و نمایش سریع نتایج بود که در تمامی نظام‌ها به‌صورت مشترک وجود داشته است. ارائه مجموعه‌ای از مفاهیم در حوزه‌ای خاص و توسعه اصطلاح‌نامه‌ای نیز در میان ویژگی‌های جست‌وجوی مفهومی در بیشتر نظام‌ها به کار رفته بود. در جست‌وجوی مفهومی «پایگاه اطلاعاتی اریک» با ۲۲ امتیاز در بالاترین سطح امتیازات قرار گرفت و این پایگاه از نظر جست‌وجوی مفهومی ویژگی‌های بیشتری را در خود نشان داد.

پیشنهاد برای جست‌وجو شامل ارائه پیشنهاد بر اساس اصطلاحات نتایج مرتبط، ارائه پیشنهاد بر اساس اصطلاح‌نامه، پیشنهاد رفع ابهام اصطلاحات پرس‌وجو بود و در بررسی نظام‌ها این ویژگی‌ها دیده نشد. تنها مورد موجود ارائه پیشنهاد بر اساس اصطلاحات نتایج مرتبط بود که در «کتابخانه ملی استرالیا» و در بخشی با عنوان «Similar Items»، اسناد مشابه با نتایج را نشان می‌داد و در «پایگاه اطلاعاتی نورمگز» نیز قسمتی با عنوان پیشنهاد برای جست‌وجو وجود دارد که بر اساس اصطلاحات موجود در اصطلاح‌نامه کلماتی برای جست‌وجو پیشنهاد می‌کند. پیشنهاد رفع ابهام اصطلاحات پرس‌وجو به‌عنوان ویژگی مهمی که در نظام‌های معنایی به کار می‌رود، در هیچ‌یک از این نظام‌ها مشاهده نشد. «کتابخانه ملی استرالیا» و «پایگاه مجلات نورمگز» تنها مواردی بودند که پیشنهاد برای جست‌وجو را ارائه می‌کردند.

جست‌وجوی مترادف‌ها شامل جست‌وجوی مترادف‌ها با اصطلاح‌نامه، جست‌وجوی مترادف‌ها با هستی‌شناسی است. ویژگی استفاده از اصطلاح‌نامه‌ها در جست‌وجوی مترادف‌ها در نیمی از نظام‌ها به کار رفته است و این موضوع اهمیت حضور آن را در ویژگی‌های معنایی نشان می‌دهد. در میان نظام‌های بررسی شده جست‌وجوی مترادف‌ها در سه نظام «نورمگز»، «کتابخانه کنگره آمریکا» و «اریک» انجام می‌شد.

استفاده از هستی‌شناسی به‌عنوان یکی دیگر از ویژگی‌های معنایی مورد توجه شامل استفاده از رابط کاربر برای تایپ اصطلاحات در هستی‌شناسی، استفاده از هستی‌شناسی برای توسعه پرس‌وجو، ارائه ساختار فنی هستی‌شناسی در رابط کاربر، حمایت از تاکتیک‌های جست‌وجو شامل تاکتیک‌های CONTRARY، RELATE، SUB، SUPER، از این ویژگی‌ها XHAUST، SPECIFY، PINPOINT، PARALLEL، SELECT، RECORD در نظام‌های مورد بررسی استفاده نشده است و از میان تاکتیک‌هایی که ویژگی‌های معنایی را حمایت می‌کنند تنها تاکتیک‌های SUB و SUPER، که به ترتیب حرکت سلسله‌مراتبی به اصطلاح اعم (بالا‌تر) و حرکت سلسله‌مراتبی به اصطلاح اخص (فرعی) است، مورد توجه این نظام‌ها قرار گرفتند. این تاکتیک‌ها اشاره به روابط اصطلاح‌نامه‌ای اعم و اخص دارد و نشان می‌دهد که این نظام‌ها تاکنون توانسته‌اند روابط معنایی موجود در اصطلاح‌نامه‌ها را در بازبازی اطلاعات به کار گیرند. «کتابخانه ملی استرالیا» و «پایگاه اطلاعاتی اریک» از نظر توجه به ویژگی‌های هستی‌شناسی در سطح بالاتری نسبت به سایر نظام‌ها قرار داشتند.

ویژگی‌های مربوط به نمایش ویژگی‌های معنایی شامل ارائه بخش مرتبط با نتایج به شکل ساده، نمایش موضوعات مرتبط، نمایش برجسته کلیدواژه‌های مرتبط به با پرس‌وجو، تعمیم یک مفهوم کلی به اجزای آن، ارائه مدلی تصویری از نتایج کاوش، نمایش رکورد به همراه خلاصه، ارائه نتایج یکسان برای پرس‌وجوهای حاوی نحو متفاوت و معنای مشابه، نمایش اشکال مختلف دستوری پرس‌وجو، و ارائه تعریف بود که ارائه بخش مرتبط با نتایج به شکل ساده و نمایش رکورد به همراه خلاصه در تمامی نظام‌ها وجود داشت. نمایش برجسته کلیدواژه‌های مرتبط به با پرس‌وجو که نمونه‌ای ساده از نمایش ویژگی‌های معنایی است، تنها در نیمی از نظام‌ها مورد توجه قرار گرفته است. به‌طور کلی، در مورد ویژگی‌های مربوط به نمایش، در نظام‌های مورد بررسی «نورمگز» و «کتابخانه ملی استرالیا» از لحاظ ویژگی‌های مربوط به نمایش در سطح بالاتری قرار گرفتند.

علاوه بر این، نتایج نشان داد که در بین نظام‌های بازیابی اطلاعات مورد بررسی، «پایگاه اطلاعاتی اریک» با ۳۲ امتیاز از مجموع ۵۶ امتیازی که امکان اکتساب آن وجود داشت، بهتر از سایر نظام‌ها در زمینه معیارهای معنایی بوده است. پس از آن «کتابخانه ملی استرالیا» با ۲۴ امتیاز رتبه دوم را دارد. این نشان می‌دهد که می‌توان تا حد زیادی از ویژگی‌های این نظام‌ها در طراحی نظام بازیابی معنایی فارسی الگو گرفت. «آرشیو ملی استرالیا»، «موزه بریتانیا» و «پایگاه اطلاعاتی نورمگز» نیز به ترتیب با ۱۳، ۱۴ و ۱۶ امتیاز از ۵۶ امتیاز را کسب کرده‌اند و این نشان از ضعف این نظام‌ها در ویژگی‌های معنایی است.

فهرست منابع

اکبری، احمد. ۱۳۸۹. پایگاه‌های مرکز تحقیقات کامپیوتری علوم اسلامی در یک نگاه. فصلنامه ره‌آورد نور. ۵۰ (۳۳): ۳۴-۴۳. <http://www.rahavardnoor.com/index.php/authors/item/155-tahghighat> (دسترسی در ۱۳۹۵/۷/۱۵).

دری، راحله. ۱۳۹۳. مقایسه و ارزیابی موتورهای جست‌وجوی معنایی. پژوهشنامه پردازش و مدیریت اطلاعات. ۴۸۷-۴۶۷ (۲): ۳۰.

شرکت خدمات کامپیوتری پیک داده‌گستر. ۱۳۸۳. استفاده از اصطلاح‌نامه در پایگاه‌های اطلاعاتی مرکز اطلاعات و مدارک علمی ایران. تهران: پژوهشگاه علوم و فناوری اطلاعات ایران.

علیمراد، مصطفی. ۱۳۹۴. نگرشی به گونه‌های جست‌وجو و بازیابی اطلاعات. فصلنامه ره‌آورد نور ۵۹ (۶۲): ۲۶.

References

- Andago, Martino, Teh P. L. Phoebe, and Bassam A. M. Thanoun. 2010. Evaluation of a Semantic Search Engine against a Keyword Search Engine Using First 20 Precision. *International Journal for the Advancement of Science & Arts* 1 (2): 55-63.
- APAIS Thesaurus. (n. d.). <https://www.nla.gov.au/apais/thesaurus/about.html>. (Accessed Dec. 24, 2016)
- Azizan, A, Z. A. Bakar, N. K. Ismail, and M. F. Amron. 2013. Interface features of semantic web search engine. *IEEE Conference on e-Learning, e-Management and e-Services*. Kucing, Malaysia.
- Bates, M. J. 1990. Where should the person stop and the information search interface start? *Information Processing & Management* 26 (5): 575-591.
- García, Elena, and Miguel-Ángel Sicilia Sicilia. 2003. User Interface Tactics in Ontology-Based Information Seeking. *PsychNology Journal* 1 (3): 242 – 255.
- Hynek, Jiř. 2002. *Document Classification in a Digital Library*. Pilsen: University of West Bohemia.
- Jones, Karen Sparck, and Martin Kay. 1973. *Linguistics and information science*. New York and London: Academic press.
- Linckels, Serge, and Christoph Meinel. 2011. *E-Librarian Service; User-Friendly Semantic Search in Digital Libraries*. Berlin Heidelberg: Springer-Verlag.
- National library of Australia*. <https://www.nla.gov.au/australian-public-affairs-information-service>. (accessed Dec. 24, 2016).
- O' leary, Mick. 2010. Hakia gets serious with semantic search. *Information today* 27 (6): 38-43.
- Read, Allen Walker. 1942. The lexicographer and general semantics. *General Semantics Monographs* (Institute of general semantics, Lakeville, Connecticut) (3).
- Styltsgiv, Henrik Bulskov. 2006. *Ontology-based Information Retrieval*. PhD thesis, computer science section Roskilde University. Denmark.
- Sudeepthi, G., G. Anuradha, and P. Babu. 2012. A Survey on Semantic Web Search Engine. *IJCSI International Journal of Computer Science Issues* 9 (2): 241-245.
- Sutcliffe, A, and M Ennis. 1998. Towards a cognitive theory of information retrieval. *Interacting with Computers* 10 (3): 321-351.
- Thesauri & Controlled Vocabularies*. 2016. <http://www.loc.gov/library/libarch-thesauri.html>. (accessed Dec. 23, 2016)
- Turner, Duygu, M. A Shah, and Yiltan Bitirim. 2009. An empirical evaluation on semanticsearch performance of keyword-based and semantic search engines: google, yahoo, msn. In proceedings of Fourth international conference on internet monitoring. Venice, Mestre.

المیرا کریمی

دارای مدرک کارشناسی ارشد رشته علم اطلاعات و دانش‌شناسی از دانشگاه تربیت مدرس تهران است. ایشان هم‌اکنون دانشجوی دکتری رشته علم اطلاعات و دانش‌شناسی پژوهشگاه علوم و فناوری اطلاعات ایران است.

هستی‌شناسی و بازیابی اطلاعات از جمله علایق پژوهشی وی است.



محمود بابایی

دارای مدرک تحصیلی دکتری در رشته علوم ارتباطات از دانشگاه علامه طباطبایی است. ایشان استادیار پژوهشگاه علوم و فناوری اطلاعات ایران است. وی هم‌اکنون مدیرعامل شرکت دانش‌بنیان «سبوی دانش پارسیان» است. سایبرنتیک از جمله علایق پژوهشی وی است.



ملوک‌السادات حسینی بهشتی

دارای مدرک تحصیلی دکتری زبان‌شناسی همگانی از دانشگاه تهران است. ایشان هم‌اکنون استادیار و رئیس پژوهشکده مدیریت دانش پژوهشگاه علوم و فناوری اطلاعات ایران است.

